Instr Sci (2010) 38:173–208
DOI 10.1007/s11251-008-9083-4

Graphical means for inspecting qualitative models of system behaviour

Anders Bouwer · Bert Bredeweg

Received: 21 August 2007 / Accepted: 19 November 2008 / Published online: 20 December 2008

Abstract This article presents the design and evaluation of a tool for inspecting conceptual models of system behaviour. The basis for this research is the GARP framework for qualitative simulation. This framework includes modelling primitives, such as entities, quantities and causal dependencies, which are combined into model fragments and scenarios. Given a library of model fragments and a scenario describing an initial situation, the qualitative simulation engine generates predictions in the form of a state–transition graph. This rich knowledge representation has potential for educational purposes. However, communicating the contents of simulation models effectively to learners is not trivial. The predicate logic format used by GARP is not easy for non-experts to understand, and a simulation often contains so much information that it is difficult to get an overview while still having access to detailed information. To address these problems, a tool has been developed that generates graphical representations of the information contained in a qualitative simulation. This tool, named VISIGARP, incorporates a vocabulary of graphical elements for model ingredients and relationships, and combines these into interactive diagrams. VISIGARP has been evaluated by thirty students, with promising results, using a setup which included simulation results and exercises about Brazilian Cerrado ecology.

Keywords Qualitative simulation · Qualitative reasoning and modelling · Conceptual models of system behaviour · Diagrammatic visualization · Simulation-based learning environments

A. Bouwer (✉) · B. Bredeweg
Human-Computer Studies Laboratory, Faculty of Science, University of Amsterdam, Kruislaan 419, Matrix I, 1098 VA Amsterdam, The Netherlands
e-mail: A.J.Bouwer@uva.nl

B. Bredeweg
e-mail: B.Bredeweg@uva.nl

Introduction

The development of computer-based software artefacts that support learners in understanding systems and their behaviour is an active area of research (e.g. see Forbus and Feltovich 2001; van Joolingen et al. 2007). This article discusses the construction and evaluation of VisiGarp, a domain-independent model inspection tool that enables learners to acquire conceptual knowledge about the behaviour of systems.

Research on science teaching has pointed out the importance of learners constructing conceptual interpretations of system behaviour (Mettes and Roossink 1981; Elio and Sharf 1990; Ploetzner and Spada 1998; Frederiksen and White 2002). Conceptual analysis of system behaviour entails several subtasks, which serve different purposes: system identification, explication of assumptions, recognition of processes, identification of distinct behavioural states, evaluation and explanation. System identification is to distinguish between relevant and irrelevant aspects in a problem situation, as well as to distinguish between the structural and behavioural aspects. The recognition of processes relates to the behavioural aspects involving causality: how do quantities affect each other? The explication of assumptions is necessary to determine which laws and equations (numerical, or qualititative) can be applied to a certain problem situation. When processes lead to changes in the situation, it is useful to identify distinct behavioural states, based on important landmark values for quantities. Often, the qualitative analysis precedes and informs the process of solving the problem numerically, and afterwards, a domain expert also typically uses a qualitative interpretation to evaluate, explain and defend the solution to the problem (de Kleer 1990). Qualitative solutions can also be considered as goals in their own right, as is the case in many educational situations, especially in pre-college education (Jackson et al. 1996). In summary, constructing a conceptual interpretation is an important means for understanding and explaining system behaviour.

Qualitative Reasoning (QR) is an area of artificial intelligence that develops theories and implements software artefacts that automate conceptual analyses, as discussed above, using computers (Weld and de Kleer 1990; Bredeweg and Struss 2003). Several QR engines exist that can generate qualitative simulations of system behaviour based on a detailed model, which may include the structure of the system, starting conditions and assumptions and knowledge about processes (e.g. Envision—de Kleer and Brown 1984; Qpe—Forbus 1984; Garp—Bredeweg 1992; Qsim—Kuipers 1994). However, it is not straightforward to use such tools in educational settings. Qualitative models are usually large[1] and they exist mainly as propositions in programming languages such as Lisp and Prolog, which are difficult for non-programmers to interpret.

The aim of our research is to make the knowledge contained in qualitative simulations accessible and insightful for learners. We do this by creating and investigating interactive tools that allow learners to manipulate and inspect qualitative simulations. In our research, we use the Garp framework for qualitative simulation, because it combines a rich knowledge representation with facilities for multi-state simulation (Bredeweg et al. 2006a). We take a graphical approach to facilitate the interaction, using interactive diagrams embedded in a graphical user interface. Diagrammatic representations are powerful means for communicating information, for various reasons. There is a closer, more analogical, mapping possible between diagrams and the real world than is possible with propositional representations, such as text or logic statements (Kulpa 1994). Because there is no need to present information sequentially, diagrams allow a more flexible and

[1] An average model can easily contain a thousand statements (de Koning et al. 2000).

meaningful use of space (Tversky 1995). Diagrams can easily be extended to incorporate or derive extra information, e.g. by drawing relationships or making annotations (Norman 1993; Suthers 2003). Therefore, we expect a benefit from moving from the native Garp representation (predicate logic statements in Prolog) to a graphical user interface with interactive diagrammatic representations.

But what should these diagrammatic representations look like and how should they be embedded in a graphical user interface? When looking at existing simulation-based learning environments that use graphical user interfaces, such as Steamer (Hollan et al. 1984) and CyclePad (Forbus et al. 1999), it is immediately apparent that they are tailored to a specific domain, with pictorial elements which show prototypical elements in that domain. In the case of Steamer, which simulates a power plant, the user looks at a screen of gauges which reflect the state of components in the simulation; in CyclePad, the user can graphically assemble a complex thermodynamics system from a selection of parts, represented by pictorial icons and connections between them. However, Garp is a domain-independent simulation framework, and adopting a particular set of pictorial elements is therefore not viable. Instead, a more generic approach is required.

To address the questions and concerns raised above, we present a tool for inspecting qualitative simulations, called VisiGarp, based on a new domain-independent diagrammatic vocabulary. VisiGarp generates several types of diagrams, which offer different views of the information contained in a qualitative simulation. To evaluate the principles behind its design, and to see to what extent people can learn from qualitative simulation models, we have carried out an evaluation study with thirty students working with VisiGarp.

The contents of this article are as follows. The second section describes qualitative reasoning and modelling (QRM), with a special focus on the Garp framework. The third section presents the design of VisiGarp, with a focus on the approach for generating interactive diagrams. The fourth section describes the domain of Brazilian Cerrado ecology as an example application of qualitative modelling. The fifth section shows how VisiGarp can be used for running and inspecting qualitative simulations, illustrated throughout with examples using the qualitative model of the Brazilian Cerrado domain. The sixth section describes the materials and setup for the evaluation study with VisiGarp. The seventh section presents the results of this study. The eighth section contains a general discussion that considers the findings in the context of related work, and future research. The last section concludes this article.

Qualitative reasoning and modelling

In many domains, if not all, it is desirable to understand the behaviour of a particular system to such a degree that it can be predicted. The field of QRM aims to develop models of systems which capture this level of conceptual understanding (Weld and de Kleer 1990). When this conceptual knowledge is made explicit, so that it can be used to explain the results of simulations, we can speak of articulate simulation models (Forbus 1988; Bredeweg and Winkels 1998). This requires knowledge of various kinds to be represented.

First of all, it is necessary to describe the structure of the system to be studied in terms of its components: what are the main entities, what are the structural relationships between those entities, what other features should be included in the model? The structure will determine the focus and scope of the model, and represents the aspects of the system that are, in principle, stable. The behaviour, on the other hand, refers to variables, or *quantities*, that belong to certain entities in the model. In qualitative models, the value of a quantity is represented as being at a distinctive point value (*landmark*), or in an interval between, above, or below such points. Based on these qualitative distinctions, qualitative states can be distinguished, which represent qualitatively different states of behaviour (Kuipers 1994). These states are connected by transitions, so that paths can be recognized which specify possible behaviours of the system. Together, they constitute the state–transition graph.

To arrive at these behaviours, processes must be identified that affect quantities in the model. To determine when processes are active and when they are not, it is necessary to specify the assumptions and conditions under which each process applies. Rather than using differential equations as in many numerical modelling methods, qualitative modelling uses the notions of influences and propagation of these influences, both of which can be positive or negative (Forbus 1984). These causal dependencies are directed to specify the direction of causality, which can be used for explanation purposes. This causal knowledge determines whether a quantity will increase, decrease, or stay steady. When a quantity is changing, it will move towards its next landmark value. If the description of the current situation does not uniquely determine what will happen, multiple alternative behaviours are predicted, which can result in branches in the state–transition graph. In different branches, quantities will reach their next landmark or not, or in a different order. This corresponds to our common sense knowledge that different things can happen given different conditions and assumptions.

Qualitative reasoning according to Garp

In order to generate predictions based on a qualitative model, it is necessary to adhere to principled rules about how to derive behaviour from structure. In the Garp framework for QRM (Bredeweg 1992), these rules are implemented in a simulation engine. Given a scenario (a description of a particular situation) and a library of model fragments (the generic domain knowledge) as input, the simulation engine produces a state–transition graph as output, consisting of states and state–transitions.

A scenario consists of a description of the system structure, in terms of its entities, attributes and structural relationships, together with initial values for the relevant quantities of the system. A model fragment represents a chunk of generic knowledge about a particular topic. There are model fragments in different categories: *description views*, which describe situations, *processes and agents*, which introduce causal influences to the system, *assumptions*, which represent assumptions about the behaviour of systems, and *composition views*, which combine multiple model fragments to represent more complex concepts. All model fragments represent generic knowledge that the simulation engine will try to apply to a given situation. Each model fragment consists of conditions and consequences. The conditions may include structural constraints (e.g. specifying the type of entity, attributes and structural relationships that must be present), as well as behavioural constraints (e.g. specifying that a quantity value must be greater than some threshold value in its quantity space, or greater than some other quantity). A model fragment is applicable to a state in the simulation if its conditions are true, in which case its consequences will be added to that state. The consequences may introduce quantities that become relevant under certain conditions, as well as dependencies (e.g. influences, in the case of process model fragments).

During the simulation process, the simulation engine applies model fragments that match the scenario to infer knowledge about the possible state(s) that the system could be in. If processes are active, their influences can cause quantities to increase or decrease. This may lead to transitions to successor states, if a quantity reaches a value that is qualitatively different from the previous state. For all successor states, the process of applying model fragments is carried out recursively, until no more successor states can be found and the simulation is complete.

Both the input required by Garp and the output it generates are represented in predicate logic format (Prolog). This allows the computer to generate a simulation by determining logical consequences from the input. Unfortunately, this predicate logic format is challenging for humans to read and, with models of reasonable size, it is difficult to get an overview of the information, especially for novices. Therefore, tools are necessary to facilitate interpretation of the results of qualitative simulations and their underlying models. This has been the motivation behind the design of VisiGarp.

The design of VisiGarp

We have designed a tool, named VisiGarp (implemented in Swi-Prolog/Xpce), to provide users (i.e. students) with automatically generated diagrams to interactively access information from Garp simulations.[2] An overview of Garp and VisiGarp and their relationship is presented in Fig. 1.

The user interacts with VisiGarp to select, control and inspect qualitative simulations generated by Garp. To this end, VisiGarp generates interactive diagrams of various kinds, visualizing the knowledge contained in a simulation. The user can select a scenario to start a new simulation, or open a saved simulation. With the simulation controls, the Garp simulation engine can be controlled in various ways, to run a simulation completely, or constrain the simulation process in the direction of particular states. The work described in this article focuses on how users can inspect simulations.

Generating interactive diagrams

The knowledge contained in qualitative simulations is highly structured. This structure can be exploited to generate appropriate visualizations, surpassing the possibilities of linear text. Graphical representations, such as block diagrams, trees and graphs can facilitate search, recognition and inference processes (Larkin and Simon 1987; Norman 1993). Diagrammatic representations can support the process of relating and comparing different information elements by making conceptual information more directly available for perception (Kulpa 1994). Good visualizations make structural aspects of knowledge explicit, which can facilitate internalization of complex concepts (Cheng et al. 2001).

To take advantage of this potential, we have developed a visual language for representing information from qualitative simulations in the Garp framework. This visual language consists of visual elements, which are combined in different ways to form particular visualizations (or *views*) supporting specific reasoning tasks. To inform the design of these visualizations, we have laid down the following guiding principles:

[2] See http://www.garp3.org and http://www.swi-prolog.org for more information on Garp and Swi-Prolog.

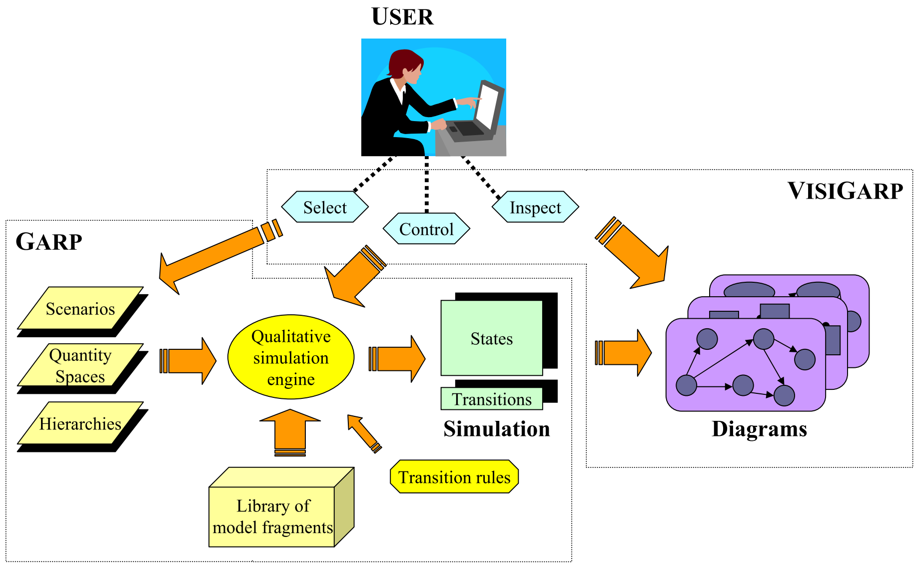


Fig. 1 An overview of GARP and VISIGARP

- Information related to a specific entity is displayed within the visual element representing that entity, using the container-metaphor as much as possible.
- A relationship between entities is displayed as a separate visual element connecting the entities involved, as much as possible.
- If an information element has more detailed information associated with it, it should be displayed as a separate visual element, large enough to be selected and to incorporate connections.
- Information which is only meaningful in a particular context should be displayed within that context, highlighted if necessary.
- If an information element is shown in a context in which further detail is not necessary, a simple textual label will suffice.
- When multiple types of relationships are combined in one view, labelled links are used to distinguish them; if all relationships to be displayed in a view are of the same type, unlabelled links, or indentation (for relationships between textual elements) will suffice.
- When multiple types of information are combined in one view, options are provided to turn on or off their display.

The visual primitives in the language include circular, rectangular and oval shapes of variable sizes for different kinds of entities, and lines, arrows, inclusion, ordering and indentation for different kinds of relations. Note that our visualization approach does not require any domain-specific knowledge, symbols, or complex shapes. This is done intentionally to ensure flexibility for use in multiple domains, and to create a graphical language that can also be used when drawing by hand.

Figure 2 presents the mapping from the main GARP modelling primitives to the visual primitives designed for VISIGARP. Where necessary, the visual primitives are shown within their visual context. Each of the mappings is described below.

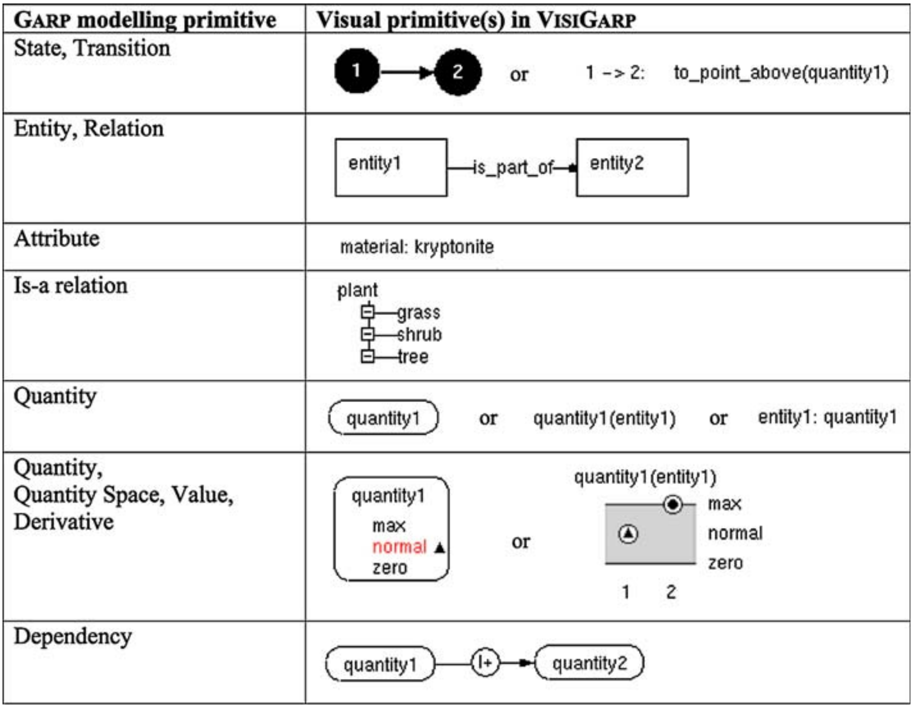


Fig. 2 The mapping from GARP modelling primitives to their corresponding visual primitives in VISIGARP

As shown in Fig. 2, a state is displayed as a coloured circle (displayed in this article as greyscales) labelled with an identification number, or it is referred to by the number only. A transition is displayed as an arrow from the originating state to the resulting state, in graphics, or text format. In the latter case, it is followed by the name of the transition rule that resulted in the transition. An entity is displayed as a labelled box, or just by the label itself. A relation is displayed as a labelled line between two entities, with a dot indicating the reading direction. In textual format, the line and the dot are left out. An attribute with its attribute value is displayed as ‘Attribute: Value’. An is-a relation between an entity type (e.g. grass, shrub, tree) and its supertype (e.g. plant) is displayed graphically by indentation, with small lines and boxes to connect the types. A quantity is displayed graphically as a labelled box with round edges. In text format, a quantity is denoted by its name, which may be followed or preceded by the name of the entity it belongs to. A quantity space is shown by vertically listing the values it contains, ordered from low values at the bottom to high values at the top. A quantity space is shown either within the quantity it belongs to, or separately, with the value labels next to a graphical space (e.g. for showing a value history). In the latter case, the difference between points and intervals is shown as lines and grey areas (i.e. zero and max are points, and normal is an interval value). A quantity value is displayed graphically as a highlighted value label in the context of a quantity space (e.g. quantity1 initially has the value *normal*), or as a small white circle on a line or between lines within a quantity space (in the case of the value history). A quantity derivative is displayed graphically as a small black triangle pointing upwards (plus) or downwards (min), or a small black circle (zero). This is shown besides the value label, or within the value circle. A dependency is shown graphically where possible as a labelled arrow (or line, in the case of an undirected dependency) between two quantities or values.

Composition and meaningful space

To show learners how certain model concepts structurally fit together, visual primitives are combined graphically by positioning them besides, or inside each other. For example, Fig. 3 displays two entities, both of which have a quantity associated to them. Entity1 is made of the material kryptonite, and is part of Entity2. Quantity1 has the value normal and is steady, while Quantity2 is zero and increasing. Furthermore, Quantity1 has a positive influence on Quantity2 ($I+$). In Fig. 3, all attributes and quantities belonging to the same entity are grouped together *within* the block representing that entity. Furthermore, within a quantity node, the quantity space is displayed as a vertical list of all possible values. The current value in the selected state is highlighted in red (displayed in this article as grey), and flanked by a derivative symbol indicating increase, steadiness, or decrease.

The collection of nodes, subnodes and the various kinds of links forms a kind of hierarchical graph (Harel 1995). This facilitates recognition of dependencies within subsystems, and dependencies crossing subsystem borders. In VisiGarp, graph edges occur only between nodes of the same type, i.e. from entity to entity, from quantity to quantity, or from value to value (not shown in the figure).

Besides the shape of a visual element, the position of a visual element with respect to its graphical context may also be used to communicate information—this is known as the use of meaningful space (Engelhardt 2002). In Fig. 3, the positioning of the highlighted value and derivative symbol on the vertical dimension is used to express information about relative order and direction of change.

Figure 4 shows how paths of behaviour are visualized in the *state–transition graph* and the *Value history view*. In the *state–transition graph* (Fig. 4a), the position of states is adjustable by the user, and a path, which may therefore have any polygonal shape, is recognizable by following the transition arrows. In the *Quantity value history view* (Fig. 4b), however, the path is straightened out and simply represented as a horizontal linear sequence of state numbers, which alleviates the need for transition arrows and leaves room for the values and derivatives to be displayed (Quantity2 increases from zero until it becomes steady in max). Here, the value of a quantity in a certain state is represented by a small circle rather than a highlighted value label, to save space and avoid repetition of labels. This circle is vertically aligned to the corresponding value label and horizontally to

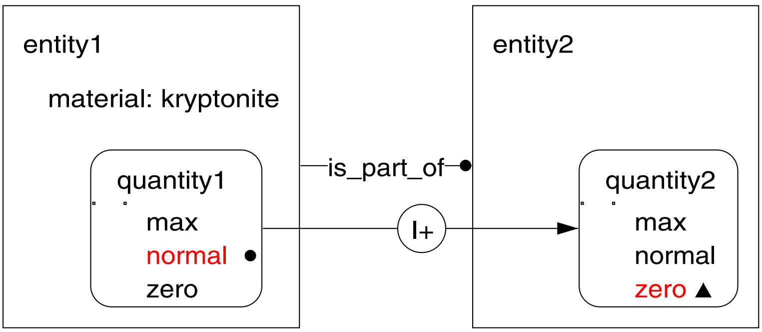


Fig. 3 Entities, attributes, quantities, quantity spaces, values, derivatives and dependencies combined

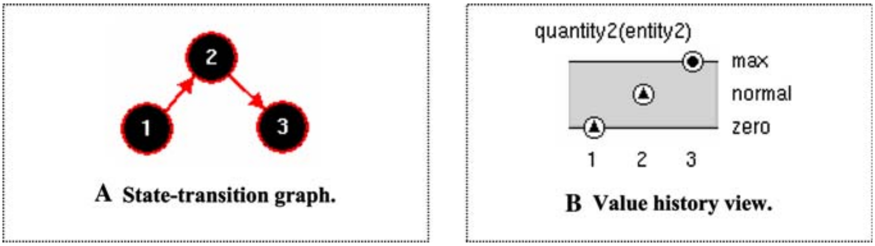


Fig. 4 Visualization of paths in the state–transition graph and the value history view

the corresponding state, on a grid with qualitative (rather than metric) dimensions. The derivative symbol is displayed inside the value circle (unless the derivative is unknown). Since this view explicitly focuses on quantity values, the representation of the quantity space includes all available detail. Point values are indicated by lines in the quantity space. Interval values do not have a line, to make clear that the value can be anywhere above/below/in between the point values which border on the interval.

Qualitative knowledge in Brazilian Cerrado ecology

In ecology, modelling and simulation are important because they complement the study of ecological systems in reality in various ways. Ecological processes, such as population dynamics, or erosion may take place at a large time scale which makes it difficult to study and experiment with them. When enough accurate numerical data is available, mathematical models can be used to analyze trends in the past, or to generate predictions. But, apart from the fact that reliable numerical data is often not available, a model that generates good predictions does not necessarily give good explanations. Besides a good fit to the data, it is considered increasingly important that models help in attaining a conceptual understanding of ecological systems and processes (Grimm 1994). The field of QR explicitly focuses on capturing such conceptual understanding. The resulting qualitative models can be used to run simulations and explain the results based on sound reasoning principles, without requiring numerical data. Building a qualitative model and running simulations of different scenarios allows exploration of alternative hypotheses and predictions, and thus aids in understanding ecological problems and potential solutions.

The example domain considered here is the ecology of the Brazilian Cerrado (Salles 1997). In this type of ecosystem, several ecological communities can occur, consisting of different plant populations of varying size (see Fig. 5). The numerous types of plants are categorized as either grass, shrubs, or trees. This allows a classification of vegetation types, ranging from Campo Limpo, consisting of mainly grassland, to Cerradão, a vegetational state with mainly trees and shrubs, and (almost) no grass. The Cerrado Succession Hypothesis suggests that a Campo Limpo may develop towards a Cerradão under the right circumstances, in terms of natural conditions and human actions, such as fire management (Pivello 1992).

The Cerrado domain has been modelled qualitatively using the Garp framework (Salles and Bredeweg 1997, 2003b). The goal of the model is to provide insight in the causal mechanisms involved in the development and management of the Cerrado ecosystem. This is considered especially useful for ecology students and future decision makers in the field of sustainable development (Bredeweg et al. 2006b).

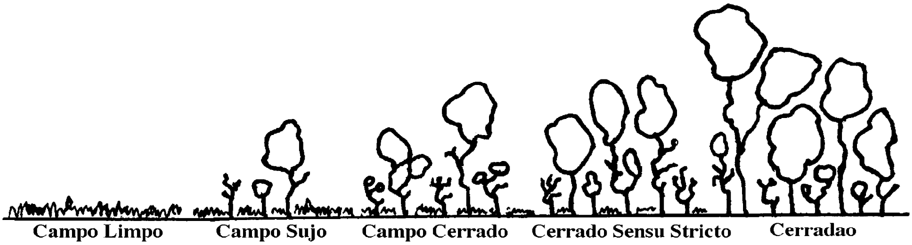


Fig. 5 Ecological communities occurring in the Cerrado ecosystem

Running and inspecting qualitative simulations

The following subsections describe the kinds of knowledge involved in qualitative simulations, and illustrate the different types of diagrams generated by VisiGarp, using the domain of Brazilian Cerrado ecology as an example.

Scenarios

A scenario describes typical aspects of the system to be simulated, including its structure and knowledge about its 'current' state, for example, the value of some of its quantities. In the Cerrado model, scenarios are defined for simple situations involving only one plant population, for intermediate situations involving pairs of interacting populations, and for complex situations, involving a complete Cerrado ecosystem, consisting of three populations: trees, shrubs and grass. For each of these, there are multiple versions with different initial values (e.g. population sizes starting at zero or max), different assumptions (i.e. open or closed populations) and different sets of processes to consider. The most complex scenarios model the Cerrado Succession Hypothesis (Pivello 1992; Salles 1997), the more simple scenarios illustrate basic principles and interaction effects without the complexity of the whole simulation. When these are put into a sequence from simple to complex, this can be regarded a curriculum of knowledge about the Brazilian Cerrado ecology (Salles and Bredeweg 2001; Salles et al. 2003).

As a running example, we consider a scenario based on the system structure presented in Fig. 6. This contains the Cerrado (shown on the top line of the figure, as a node labelled Cerrado1), which consists of a tree population, a shrub population, and a grass population (all shown one level below Cerrado1). In addition, several assumptions are associated to the populations and the Cerrado (connected by the links labelled *assumption*). For example, all three populations are assumed to be *open* populations (i.e. emigration and immigration are considered possible), and the number of emigrations is assumed to be equal to the number of immigrations. Furthermore, there is a manager, who is manager of the Cerrado and performs the task of fire control, which is set to decrease the frequency of fires in the Cerrado (shown on the top line, right of the Cerrado).

Ecological model fragments

As explained in Section "Qualitative reasoning and modelling", model fragments contain generic knowledge about system structure or behaviour, and exist in different categories: description views, assumption model fragments, process and agent model fragments and composition views.

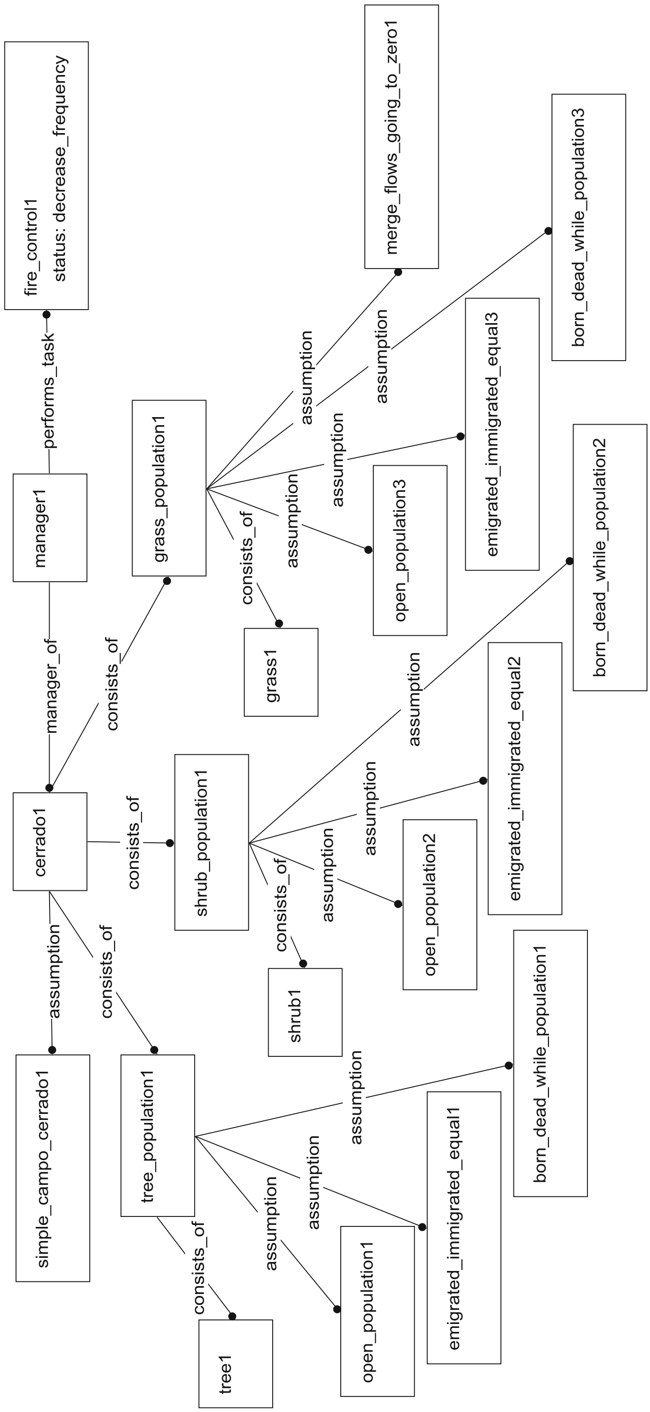


Fig. 6 The structure of the Cerrado as modelled in GARP

The library for the Cerrado domain contains 79 model fragments, each representing a particular topic that should be mastered by students learning about Cerrado ecology to enable them to adequately explain the phenomenon and make proper predictions of the system behaviour. The description views introduce concepts such as population, existing versus non-existing population, steady, increasing or decreasing population and populations of various sizes. In addition, there are description views which describe how to determine qualitative values for certain quantities in the model. Assumption model fragments include assumptions such as assume steady immigration flow, assume open (or closed) population and assume immigrated is equal to emigrated. Process model fragments represent the effects of basic ecological processes such as natality, mortality, emigration, immigration and colonization and population growth as an aggregated measure. In addition, there are Agent model fragments representing human activity, which can either decrease, or increase the fire frequency. Composition views are used to represent interaction effects between multiple populations, such as different kinds of predator–prey interactions, and amensalism (where one species affects another without being affected itself). In addition, there are composition views to represent vegetational states in the Cerrado, classified by the ratio of trees, shrubs and grass. When model fragments apply to a particular scenario, their contents will occur in the simulation results.

The state–transition graph

The state–transition graph, or behaviour graph, represents all possible states that can be predicted based on a particular scenario. States are connected sequentially by transitions, which specify the changes in values or inequalities that arise between the states involved. Figure 7 shows the state–transition graph for a simulated scenario about the Cerrado Succession Hypothesis. The figure shows a full simulation. In Fig. 7, a path has been selected between states 1 and 18, consisting of states 1, 4, 5, 12, 13, 15, and 18. This path (and indeed any path) through the simulation corresponds to a possible behaviour representing what may happen to the Cerrado.

In VisiGarp, the state–transition graph is presented in the main screen, from which all other views can be accessed. As such, it functions as a kind of overview of the simulation. States, terminations and transitions can be investigated in more detail by selecting them and clicking one of the other view buttons. If one is interested in a particular behaviour of the system represented by a path of states, the Select-mode can be set to Path instead of States (the default). When two or more states are selected in Path-mode, the shortest path through these states (if one exists) is automatically selected. This supports investigation of behavioural aspects of the Cerrado with views such as the quantity value history view.

The quantity value history view

The quantity value history view shows the values of selected quantities over the course of a simulation, or a specific behaviour path, as shown in Fig. 8. The figure shows what happens to three of the main quantities (there are 32 quantities in total) in the Cerrado simulation, following the selected path from 1 to 18. The quantity *number_of1* (the size of the tree population) starts increasing in state 4 and increases all the way to the value *max*. The quantity *number_of2* (the size of the shrub population) also starts increasing in state 4, and reaches the interval *high*. The quantity *number_of3* (the size of the grass population) decreases, from the value *max* in state 1 to *zero* in state 15, and it starts increasing again in state 18. This corresponds to what may happen to the vegetation according to the Cerrado

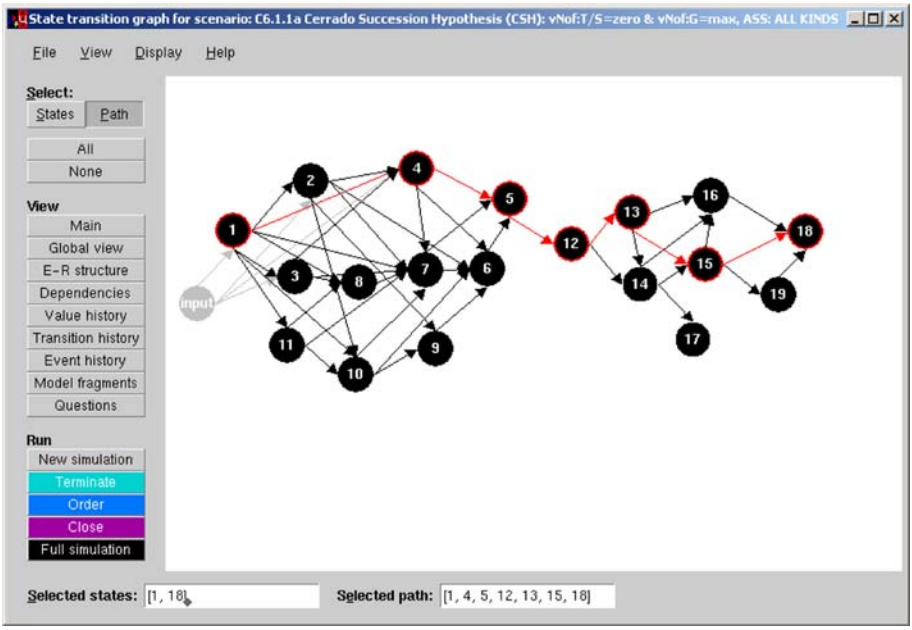


Fig. 7 state–transition graph for a simulated scenario about the Cerrado Succession Hypothesis

Succession Hypothesis: an initial state with a maximum amount of grass and no shrubs or trees (Campo Limpo) can develop into a state with no grass and a high or maximum amount of shrubs and trees (Climax-Cerradão).

The causal model

After recognizing how quantities change in the simulation, it is possible to explore the reasons behind these changes in the causal model of a particular state. The causal model consists of all quantities and the causal dependencies (influences and proportionalities) between them.

In VisiGarp, the causal model is incorporated into the Dependencies view (see Fig. 9), which allows investigation of the causes and effects of changes to quantities in the simulation. This view also shows the relevant entities (and optionally, the structural relationships) from the Entity-relations view (as shown in Fig. 6) to relate the quantities and dependencies to the system structure, and also to enforce organization in the layout. The layout can also be adjusted by hand, by dragging entities or quantities. Toggle-buttons are supplied alongside the diagram which can be used to show or hide specific types of information, such as values, derivatives, or specific types of dependencies.

Arrows labelled $I+$ / $I-$ represent direct influences, e.g. $X \xrightarrow{I+} Y$ means that a positive value of X results in an increase of Y; the arrows labelled $P+$ / $P-$ represent proportional dependencies, e.g. $Y \xrightarrow{P+} Z$ means that an increase in Y leads to an increase of Z. A negative sign reverses the effect.

Figure 9 shows the causal dependencies between the quantities of the Cerrado, the tree population, the grass population and the manager. In both populations, the four elementary processes birth, death, immigration and emigration can be recognized, affecting the

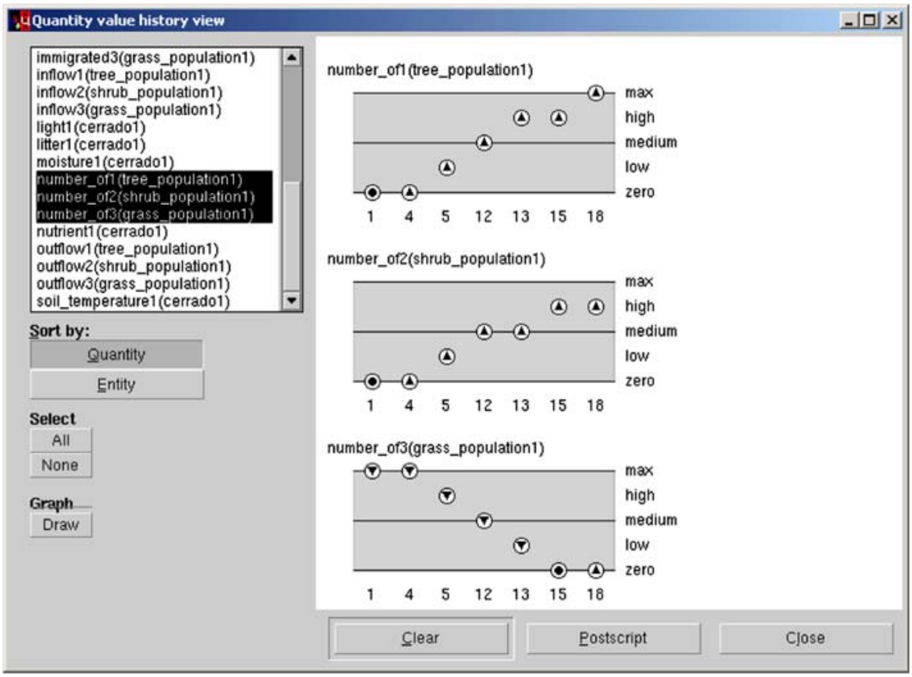


Fig. 8 Quantity value history view

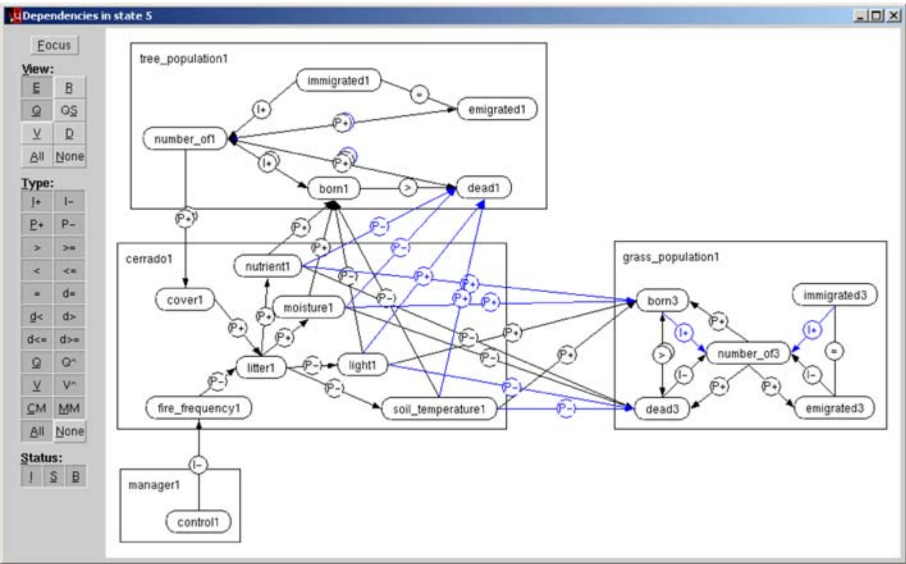


Fig. 9 Dependencies view

population size of trees (*number_of1*) and grass (*number_of3*). Within the Cerrado, the amount of litter (e.g. leaves on the ground) is affected positively by cover (the leaves on the trees) and negatively by fire frequency. However, the manager's control (which has the value plus) has a negative influence on the fire frequency, so that both causal paths lead to an increase in litter. Changes in litter propagate positively to nutrient and moisture, and negatively to light and soil temperature. These latter four factors affect the birth and death rate of the tree and grass populations. Regarding the trees, both the increase of nutrient and moisture and the decrease of soil temperature and light have a positive effect on the birth of trees (born1), via $P+$ and $P-$ dependencies, respectively. Regarding the grass, the effects are different, however. Here, the decrease of soil temperature and light have a negative effect (via $P+$ dependencies) on the birth of grass (born3). Together, these dependencies express the knowledge that more trees are 'born' when conditions are relatively cool and dark, whereas the amount of grass decreases under these conditions.

Evaluation of VisiGarp: materials and setup

VisiGarp has been fully implemented, allowing system evaluation. As the design is centred around the generation of visualizations and organizing them in a graphical user interface, the focus of the evaluation is on the following questions:

- Is VisiGarp easy to use?
- Are the visualizations understandable?

To find out to what extent students find VisiGarp easy to use, or learn to use, they are asked to work with VisiGarp to fulfil tasks concerned with the Cerrado ecology domain, with only a short introduction beforehand. To determine the extent to which the visualizations generated by VisiGarp are understandable, the students are required to fill in questionnaires with questions about the meaning of diagram elements. In addition, the students are asked to do tests about the Cerrado ecosystem, to find out whether they understood the domain material. Details about the setup of the study are found in the following subsections.

Simulations

The domain library used in this study is the Cerrado Succession Hypothesis (CSH) model described in Section "Qualitative knowledge in Brazilian Cerrado ecology". This model includes an extensive set of scenarios, including single populations involving a few processes, interactions between two populations, and more complex scenarios about the Cerrado ecosystem as a whole (Salles and Bredeweg 2001). Two scenarios were selected for the evaluation study: a simple scenario, with one plant population and the influence of four basic processes (birth, death, immigration and emigration), and a complex scenario based on the Cerrado Succession Hypothesis, which, in addition to the processes present in the simple scenario, also includes environmental factors such as soil temperature, fertility, light, cover, moisture, nutrients and fire frequency. The complexity of both simulation scenarios in terms of the number of elements is shown in Table 1.

Table 1 The complexity of simulations 1 and 2

Number of elements	Simulation 1	Simulation 2
States	8	19
Entities & relations	7	39
Quantities	8	32
Dependencies	12	127

Participants

In this study, thirty first-year psychology students participated (7 male, 23 female), who were given study credit points as a reward.

Materials

An overview of the experimental setup is shown in Table 2.

The details of each step of the procedure are described below.[3]

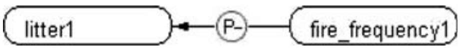
1. A short introduction was given by the experimenter. In a few minutes, the goals of VisiGarp were explained and the experimental procedure was described by reading aloud a pre-written text.
2. A paper-based pre-test of domain-specific knowledge about the Cerrado was given to the participants. To test and compensate for any possible difference in difficulty level between the pre- and post-test, two versions of the test were created, CSH-A and CSH-B. They contained equivalent questions, with some elements changed. Half of the participants received test CSH-A, the other half received test CSH-B as pre-test. An example question from this test is given below.[4]

   CSH Domain test CSH-B, question number 4.
   Does regulation of the fire frequency within the Cerrado vegetation (for example by appointing a manager who can increase or decrease the fire frequency) have an effect on the tree population?

   a. Yes, if the fire frequency is increased, the tree population will increase
   b. *Yes, if the fire frequency is increased, the tree population will decrease*
   c. No, the fire frequency and the tree population are independent of each other
   d. No, the fire frequency cannot be regulated

3. A paper-based pre-test of diagrammatic representations as shown in VisiGarp was given to the participants. Here too, there were two versions of the test, DIA-A and DIA-B. Half of the participants received test DIA-A, the other half received test DIA-B as a pre-test. An example question from the diagrams test is presented below.

   Diagrams test DIA-A, question number 7.



   What is the relationship between litter1 and fire-frequency1?

   a. *If fire-frequency1 decreases, litter1 increases*
   b. If litter1 increases, fire-frequency1 increases
   c. If litter1 decreases, fire-frequency1 increases

[3] When multiple choice questions are involved, the correct answer is displayed in italics.

[4] Except where noted, all texts in the example questions, answers and VisiGarp figures have been translated from Dutch to English. The original Dutch version is available in Tjaris (2002).

Table 2 Overview of the experimental setup

Number	Task	Time (min)
1	Short introduction	<5
2	Pre-test domain-specific knowledge	<15
3	Pre-test diagrammatic representations	<10
4	Treatment: Interaction with VisiGarp	<80
5	Post-test domain-specific knowledge	<15
6	Post-test diagrammatic representations	<10
7	Questionnaire attitudes	<15
8	Comments and questions	<10

4. The treatment: interaction with VisiGarp, using two scenarios, was carried out. A simulation scenario was loaded into VisiGarp, and the participant was asked to follow the instructions and answer the questions that were presented on the screen, in a window separate from VisiGarp. The participant could switch back and forth between the question window and the VisiGarp windows. The VisiGarp interaction session contained 27 questions in total, divided over the two scenarios.
5. A paper-based post-test of domain-specific knowledge about the Cerrado was given to the participants. Those participants who received test CSH-A as pre-test, received test CSH-B as post-test, and vice versa.
6. A paper-based post-test of diagrammatic representations as shown in VisiGarp was given to the participants. Those participants who received test DIA-A as pre-test, received test DIA-B as post-test, and vice versa.
7. A paper-based questionnaire about the subject's attitudes towards VisiGarp was delivered. The attitude questionnaire consisted of 14 questions with an answer range from 1 (negative) to 5 (positive). The complete list of questions is shown in Section "Results of the evaluation study".
8. A short time for additional comments and questions was provided. At this time, participants could ask or comment about VisiGarp, the experiment or the research in general.

Details about the treatment

The goal of the treatment questions was to familiarize the students with important aspects of the Cerrado model, addressing topics ranging from system structure to system behaviour and the underlying causal model. Answering the questions required detailed inspection of the different kinds of diagrams generated by VisiGarp. The complete list of treatment questions is presented in the appendix. The simulations used in the two scenarios were as follows:

Simulation 1 The first scenario contained one plant population with the four basic ecological processes affecting the population size: birth, death, immigration and emigration. There were 10 questions about the first scenario. An example question (with instructions for operating VisiGarp) is shown below.

Session with VisiGarp, simulation 1, question 1.
In the *state–transition graph*, select state **3**. From the **View** menu, select the option

Entities and relations.
Which of the following statements is true?

a. open_population1 consists of population1.
b. *population1 consists of biological_species1.*
c. open_population1 is an assumption about biological_species1.

Simulation 2 The second scenario included the Cerrado vegetation, with three populations (trees, shrubs and grass) and several environmental factors. There were 17 questions (number 11–27) about this scenario. Two examples are given below.

Session with VisiGarp, simulation 2, question 24.
What can be deduced about the growth of the shrub population?

a. *the shrub population increases, because immigration is equal to emigration, while born is greater than dead.*
b. the shrub population increases, because immigration is greater than emigration, while born is equal to dead.
c. the shrub population increases, because immigration is greater than emigration, AND born is greater than dead.

Session with VisiGarp, simulation 2, question 25.
The factors soil temperature, light, moisture and nutrient have an effect on born and dead of the populations in the Cerrado vegetation. Which of the following statements about the tree population is correct?

a. When litter increases, the composition of soil temperature, light, moisture and nutrient changes in such a way that for the tree population, born (growth of new trees) becomes equal to dead (dying of existing trees).
b. *When litter increases, the composition of soil temperature, light, moisture and nutrient changes in such a way that for the tree population, born (growth of new trees) becomes greater than dead (dying of existing trees).*
c. When litter increases, the composition of soil temperature, light, moisture and nutrient changes in such a way that for the tree population, born (growth of new trees) becomes smaller than dead (dying of existing trees).

When the participant answered a question (by selecting one of the options a, b, or c), VisiGarp recorded the answer and displayed the next question. For the purpose of analysis, a time-stamp was created for each answer, to give an indication of the time spent on reading, interpreting and answering each question. Participants were also asked to rate the difficulty level (on a scale from 1 = easy to 5 = difficult) of the questions after each scenario. Two help documents were supplied, one about the VisiGarp interface as a whole, and one sheet which contained a list of all types of dependencies and their meaning.

Results of the evaluation study

All participants completed the experimental procedure, except for some of the treatment questions. In the following subsections, the results are shown and briefly discussed for each type of data. In the statistical tests, we assume the data come from a normal distribution; the data did not suggest that this assumption should be discarded.

Domain knowledge pre-test and post-test

Comparing the pre-test and post-test scores, 26 out of the 30 participants improved their scores from pre-test to post-test. The three participants whose scores decreased had rather high scores on the pre-test, which may have caused a ceiling effect. Averaged over all subjects, the mean scores for the pre- and post-test are depicted in Fig. 10, together with 95% confidence intervals, indicating the error margins. As can be seen, the participant's scores increased from the pre-test ($M = 42\%$, SD $= 12.1\%$) to the corresponding post-test ($M = 67\%$, SD $= 14.2\%$, $t(29) = -6.75$, $p < 0.001$).

Recall from Section "Evaluation of VisiGarp: materials and setup" that there were two versions of the test: half of the participants received test CSH-A before CSH-B and vice versa for the other half. This allowed a comparison to examine whether there was any difference in difficulty level between test CSH-A and CSH-B. As no difference in results could be found between the results on CSH-A and CSH-B for the pre-test, nor between the results on CSH-A and CSH-B for the post-test, the data for test CSH-A and test CSH-B have been combined in the results described above.

In Fig. 11, the mean scores per question are shown for both the domain knowledge pre-test and post-test. This figure shows that on the post-test, all questions except question 7 were answered correctly more than 50% of the time, whereas on the pre-test 7 out of the 10 questions were answered correctly less than 50% of the time.

The figure also shows there is some variation between the results for the different questions. Two questions are especially interesting, in this respect. Question number 7 (*What is the effect of an increase in sunlight on the amount of births for the shrub(A)/ tree(B) population? Correct answer for A and B: decrease*) shows remarkably low scores:

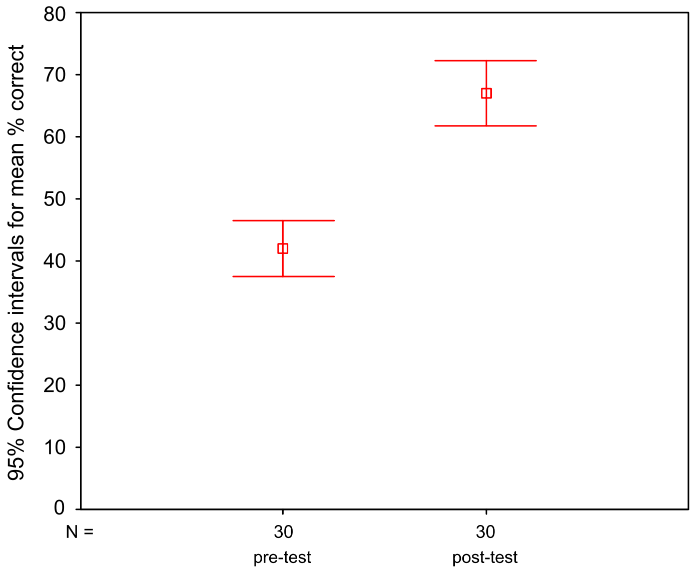


Fig. 10 Mean scores and confidence intervals for the domain knowledge pre- and post-test

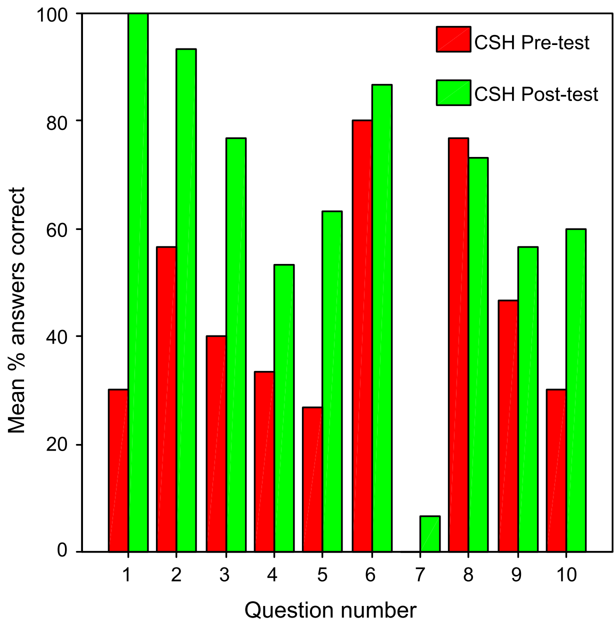


Fig. 11 Mean scores per question on the domain knowledge pre- and post-test

0% and 7% correct for the pre- and post-test, respectively. A reason for the low pre-test score may be that the participants were not aware that fewer seeds of trees and shrubs germinate when the amount of sunlight on the ground increases (in contrast to grass seeds), which causes the birth rate of trees and shrubs to drop. It is not surprising that this score did not increase very much, because the effects of sunlight on the different plant populations were not directly addressed in any of the treatment questions. The scores for question 8 are also interesting because they show a small decline. Question number 8 was, in both version A and B the same (*What is the effect of an increase in moisture of the soil on the amount of births for the grass population? Correct answer: increase*), but answering this question involved further investigation of VisiGarp. The correct answer could be found using VisiGarp to show that there is a $P+$ dependency from moisture1 to born3 of the grass_population, but because this dependency is submissive to other effects, born3 actually *decreases* in most states of the simulation. This may have caused some confusion, as the status of submissive effects was not visible in VisiGarp. Apart from question 8, the mean post-test scores were higher than the pre-test for all questions. This indicates that the increase from pre- to post-test scores is a general effect, and not localized to a few questions.

Diagram pre- and post-test

The two versions of the diagrams tests, DIA-A and DIA-B, were compared, but no significant difference was found, indicating that they had similar difficulty levels. With the results of DIA-A and DIA-B combined, the mean scores for the diagram pre- and post-test are shown in Fig. 12.

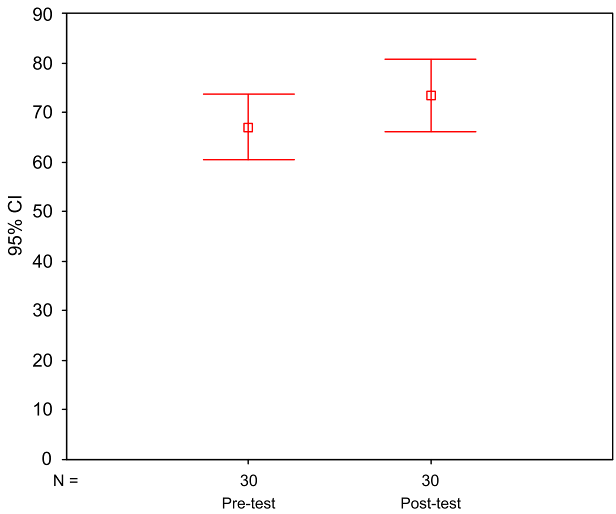


Fig. 12 Mean scores and confidence intervals for the diagram pre- and post-test

The scores on the pre-test were already reasonably high ($M = 67\%$, SD $= 17.6\%$). There was a slight increase from pre-test to post-test ($M = 73\%$, SD $= 19.5\%$), but this difference was not significant. The data suggests that the majority of diagrams were easy to understand from the start. Investigating the scores on individual questions, however, raises some interesting points. Questions about positive relationships visualized from left to right were relatively easy to answer. Questions which involved quantity space correspondences (Q-relationships), negative relationships, or relationships which were depicted from right to left (in contrast to the reading direction) were answered incorrectly more often. All three issues seemed to improve from the pre- to the post-test, but some questions remained difficult. Two of the most difficult questions are discussed in more detail:

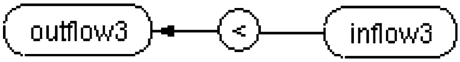
Diagrams test DIA-B, question number 2.

What is the relationship between control1 and fire-frequency1?

a. *if the value of control1 is negative, then fire-frequency1 increases*
b. if the value of control1 is negative, then fire-frequency1 decreases
c. if the value of fire-frequency1 is negative, then control1 increases

The mean score for this question on the pre-test = 43.5%, on the post-test = 60% correct. The scores on the post-test were higher, as expected, but still relatively low. The low scores here may be due to the presence of two negative elements, as double negatives are known to be difficult to understand in general. To explain the correct answer, when control1 is negative, the negative influence from control1 has a positive effect on fire-frequency1, which will therefore increase. The following question resulted in even lower scores:

Diagrams test DIA-B, question number 6.



What is the relationship between inflow3 and outflow3?

a. outflow3 is smaller than inflow3
b. *inflow3 is smaller than outflow3*
c. inflow3 is greater than outflow3

The mean score on the pre-test for this question = 13%, on the post-test = 23% correct. Again, the scores on the post-test were slightly higher, but both were very low. Here, the visualization seemed to have introduced a problem in readability. The problem stems from the similarity between the picture and the normal textual representation of the reverse relation, 'outflow < inflow'. The picture adheres to our visualization principle that text (including relationship symbols such as '<' and '>') is not rotated, while the quantity nodes can be freely moved around. This does not guarantee that the arrow follows the normal reading direction, and occasionally results in situations as in question number 6. Training may help users to follow the rules in order to read the picture correctly: the symbol '<' means 'smaller than', and the relationship holds between inflow3 and outflow3, in that direction as enforced by the arrow, so inflow3 is smaller than outflow3. To what extent this kind of training will solve the problem requires further experimentation.

Treatment exercises

The questions presented during the treatment session with VisiGarp can be found in the appendix. The first 10 questions were about scenario 1 (one population, four processes), the latter 17 dealt with scenario 2 (three populations, a manager and four processes for each population). Not all participants completed all 27 treatment questions, but everybody at least reached question 22. Only the answers given have been considered in the analysis.

The percentage of correct answers was generally high ($M = 94.3\%$, SD = 8.6%) and did not differ between scenarios 1 and 2. That the percentage of correct answers is so high is not surprising given that all answers could be found in the diagrams generated by VisiGarp, if the instructions were followed correctly. Relatively difficult questions were number 3 (the first question about a quantity space, $M = 67\%$ correct), and question number 26 (a complex question about causal and mathematical dependencies involving six interrelated quantities, $M = 70\%$).

The average time spent on each question is shown in Fig. 13. Questions 4, 6, 12, 16, 20 and 21, were all finished within 1 min, on average. Questions 1, 25 and 27 took the most time—between 4 and 5 min, on average. It is important to note that this time includes reading the question, instructions and answers, opening the right views, manipulating the layout (if necessary) and choosing an answer. Some diagrams generated by VisiGarp for simulation 2 were quite complex and cluttered, because VisiGarp places everything into view, even if that means that graphical objects will overlap. Most participants made use of the zoom buttons, 12 times per person on average over the course of the whole experiment. This made it easier to focus on a certain part of the system, e.g. the tree population, as the topic of the question at hand. Still, a lot of time was spent on adjusting the layout by dragging quantities and entities over the screen. Some students were quite good in

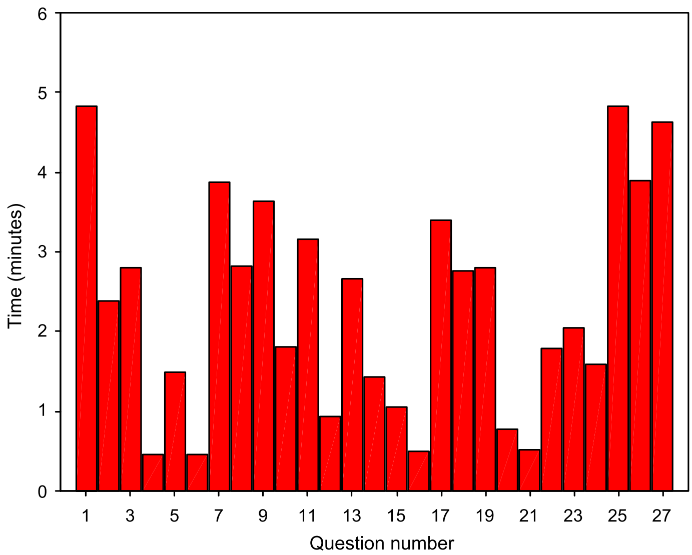


Fig. 13 Times per question in the treatment with VisiGarp

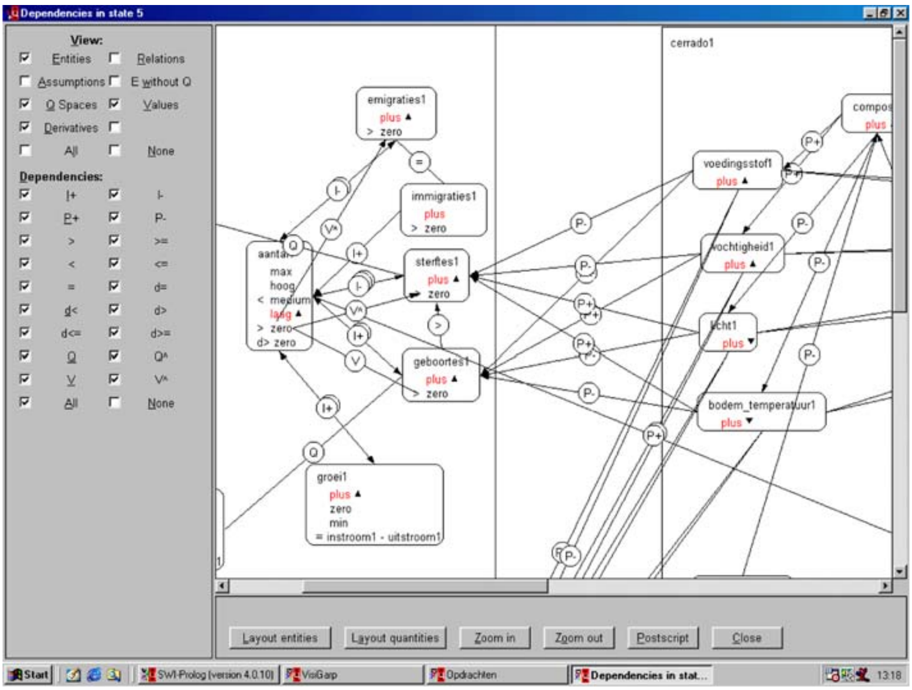


Fig. 14 A screenshot from a student working on an exercise

improving the layout this way, but others occasionally made it worse, by selecting and moving the wrong element because it partly overlapped with another. An example from a student working on an exercise is the screenshot displayed in Fig. 14. In total, the amount of time spent on all questions varied between 46 and 77 min between the fastest and slowest participant ($M = 62$ min).

The participants rated the questions in scenario 2 ($M = 3.07$) as more difficult than scenario 1 ($M = 2.53$) ($t(29) = -2.64$, $p = 0.013$). This is not suprising, given that the simulation model in scenario 2 contains much more information than in scenario 1. From their comments however, it seems that participants did not find the content difficult to understand, instead they found it challenging to find the correct answer on the screen.

Attitude questionnaire

On the scale from 1 (negative) to 5 (positive), the participants rated VisiGarp quite positively overall ($M = 3.70$, SD = 1.14), with only two participants scoring an average of <3 (neutral). Figure 15 shows the ratings per question, averaged over all subjects. For ease of reference, all questions of the attitude questionnaire are listed below, translated into English from the original version in Dutch (Tjaris 2002).

1. VisiGarp is a system which allows me to check predictions about possible changes in the Cerrado vegetation.
2. VisiGarp is a tool which helps me to explain possible changes in the Cerrado vegetation.
3. Concepts in the Cerrado domain, such as *populations, trees, mortality and emigration*, connect to the ideas I have about an ecosystem like the Cerrado.
4. The primitives used in the VisiGarp tool, such as *states* and the different *causal relations (I+, P+, etc.)* are comprehensible.
5. Because of using VisiGarp, I better understand how an ecosystem, such as the Cerrado, works.
6. The information necessary for carrying out the exercises in the experiment was easy to read from the diagrams.
7. Overall, the exercises in the experiment were easy to carry out.
8. Carrying out the exercises in the experiment has contributed to acquiring knowledge about the Cerrado domain.
9. From the different views available, I was able to find the right view very quickly.
10. The layout of the diagrams as generated by VisiGarp was well-organized.
11. The layout of the diagrams was easy to adjust.
12. The instructions on how to use VisiGarp (which preceded the exercises) have contributed to how quickly and easy I was able to find the answers to the exercises.
13. When the information did not fit onto one screen, the scroll-bar was a convenient way to see the total workspace of the different views.
14. The help document has been useful for understanding VisiGarp and finding the information desired.

The subjects considered that VisiGarp was a good tool to inspect changes in the Cerrado vegetation simulation. They were more negative about whether the layout of the diagrams generated by VisiGarp was clear (Question number 10: $M = 2.63$, SD = 1.30), but they felt the layout was easy to adjust (Question number 11: $M = 4.03$, SD = 1.00).

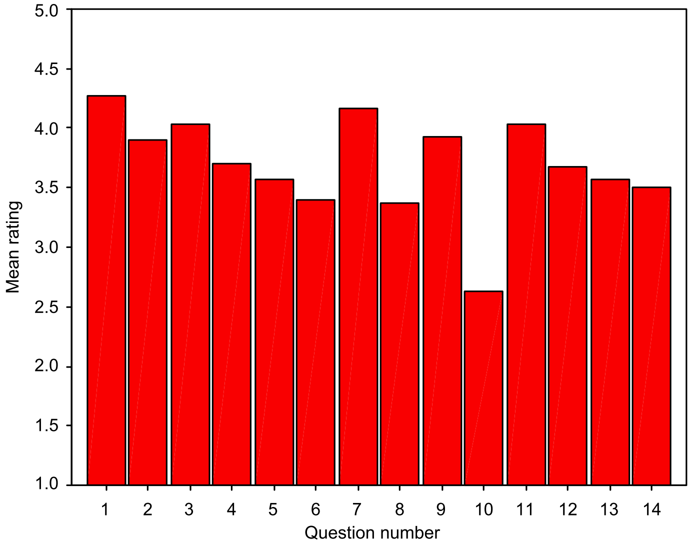


Fig. 15 Mean ratings per question

Discussion

The results of the evaluation study show that university students without any experience in QR can learn to make use of VisiGarp to complete exercises of increasing difficulty level within one and a half hours. The increase in scores from the domain pre-test to the post-test indicate that a strong learning effect occurred during the course of the experiment. This effect is remarkable, given that other experiments do not always show a clear learning effect (e.g. see van Berkum and de Jong 1991; van Joolingen 1999 for a discussion of several projects involving simulation-based learning software).

For example, in the work on SimQuest, learners interacted with numerical simulations using a discovery learning approach, but additional support was deemed necessary to create an effective learning experience (van Joolingen and de Jong 2003). In contrast, VisiGarp directly presents the conceptual knowledge to be learned, which may have led to a more focused interaction. Some caution is necessary in interpreting these results however, because no control group was used in this experiment.

It could be that the subjects learned from interacting with VisiGarp and looking at the diagrams it generates, but it is also possible that the exercise questions or the instructions were the determining factor. A third possibility is that the pre-test encouraged greater thought about the material, which led to a higher performance on the post-test. Also, it is unclear whether the learning experience would have been as efficient (or less, or even more) if the material was presented on paper instead of using VisiGarp on a computer screen. Therefore, a more comparative study is necessary to establish to what extent the different parts of VisiGarp are responsible for the learning outcome. However, as an educational tool which allows for controlling and inspecting qualitative simulation models of varying complexity, VisiGarp is unique in its flexibility, which makes it difficult to compare.

The students also rated VisiGarp quite positively afterwards, except for one issue. The layout of the diagrams generated by VisiGarp for the second, more complex simulation was often not optimal. Students could modify the layout by hand, and reported that they could do so easily, but from observations by the experimenter, it was clear that students had to spend considerable time doing this. Three possible solutions have already been implemented in later editions of the software and the accompanying user guide: improving the layout algorithm, improving the zooming mechanism and improving the user's guide. A fourth option that is being investigated, partly inspired by the work of Mallory (1998), is to aggregate the information, thereby reducing the number of graphical elements shown (Bouwer 2005).

Building and inspecting conceptual models for learning

VisiGarp is a tool for inspecting existing models, such as the Brazilian Cerrado ecology model used in the evaluation study. There are a number of tools that provide functionality for the student to build a model, in addition to inspecting existing models. These vary in terms of the type of representation used, and the kind of functionality provided.

For example, the CmapTools software (Cañas et al. 2005) is based on the research on concept maps (Novak and Gowin 1984) and has a simple representation, consisting only of concepts and relationships. With the tools, users can easily create and format their own concept map, and add comments and links from concepts in the map to external webpages, if desired. The Giant (Reichherzer et al. 1998) and Betty's Brain (Leelawong et al. 2001) are systems based around concept maps with an emphasis on the educational interaction. The main idea here is that students teach an artificial agent (or *teachable agent*), which can reason about the concept map and answer (Betty's Brain) or ask (The Giant) questions to stimulate the student to expand it. Evaluation studies with Betty's Brain suggest that constructing concept maps during studying improves learning compared to writing summaries (Leelawong et al. 2001). The Giant uses heuristic rules to derive and infer new knowledge from what the student has entered so far, but is based on a very basic knowledge representation. Betty's Brain works in a similar fashion, but also includes limited capabilities for simulation, including influence resolution to determine values and derivatives.

Model-It (Soloway et al. 1997) and Stella (developed by Richmond, 1985, see also (Doerr 1996)) are tools for creating system-dynamics models. These models include a structural diagram (Forrester 1968), that describes the causal relationships, which incorporate stocks (quantities) and flows (rates) as distinct from each other. Also included are differential equations, which drive the simulation of the system's behaviour. This makes it very different from VisiGarp, which uses a fully qualitative knowledge representation. Although the use of numerical simulations has its merits, the important advantage of qualitative simulation is that conceptually different states can be distinguished, which represent typical states of behaviour.

VModel is a tool (Forbus et al. 2001) for building qualitative process models, using the qualitative process theory, which shares some features with the Garp framework, such as the explicit representation of structural relationships and causal influences. VModel is aimed at middle-school students, a younger target audience than the high-school and university students that VisiGarp aims to address. This seems to be reflected in some of the design decisions in VModel, such as a heavier use of colours, and a smaller set of model primitives. More importantly, however, VModel only allows reasoning within a single state, and does not allow multi-state simulation of behaviour, as VisiGarp does.

An interesting difference in the visualization is that processes are represented in a similar way to entities (as black labelled boxes), with the addition of a rate parameter, which is displayed within the process box. In VisiGarp, quantities involved in processes are always associated with an entity, and processes are considered at a higher level of aggregation.

The Expound system (Mallory 1998) is based on the Qsim framework (Kuipers 1994), and is designed to explain *multi-state* simulations, like VisiGarp. Expound generates (abstracted) state–transition graphs and influence diagrams (showing quantities and causal dependencies). A useful feature of Expound which is not included in VisiGarp is the possibility of focusing on particular system parts to simplify diagrams. In a small evaluation study, the diagrams were positively rated by all three users. However, these users were qualitative reasoning experts, and no data is presented involving other types of users. Furthermore, Expound generates its diagrams as non-interactive output; they are not integrated in the interface as in VisiGarp, which limits its usefulness as a tool for end users.

To fulfil the full potential of qualitative simulation models in education (Bouwer et al. 2002), current work addresses combining model-building tools and functionality for inspecting simulations into one integrated workbench (Bredeweg et al. 2006a).

Abstract versus pictorial representations

The diagrammatic vocabulary incorporated in VisiGarp contains no pictorial representations, but a rather abstract visual language, consisting of shapes with textual labels, various kinds of connections, and several ways of using colour as a modifier. We acknowledge the value of realistic representations for the purpose of grounding, but in our research, diagrams are used to represent a conceptual view of the part of the world of interest, not what it looks like in the real world. We adhere to the view that it is the similarity in the structure, and additional affordances that makes diagrams useful (Cheng et al. 2001), rather than perceptual accuracy. Using relatively abstract representations (such as block-and-arrow diagrams) allows for more flexible use of the graphical means (such as shape, size, position and colour), by not using these means to improve on the realism of the representation, but to represent other information instead. The VisiGarp diagrams offer additional affordances in the sense that the simulation can be controlled interactively and navigated graphically, and the diagrams can be manipulated using options of the graphical interface such as dragging, selection and showing/hiding extra details. An additional advantage of this approach is that it is domain-independent. It would be interesting to investigate whether using the same visual language across different domains supports transfer of knowledge from one domain to another.

Application across domains

VisiGarp is currently in use by researchers in various domains to test and communicate simulation results of their models (Nuttle et al. 2004; Neumann and Bredeweg 2004; Tullos et al. 2004; Salles and Bredeweg 2006), and by several teachers of academic courses which incorporate qualitative modelling in ecology, physics and chemistry (Salles and Bredeweg 2003a; Alvarez-Bravo et al. 2004; Salles et al. 2004). This proves that there is a need for a tool like VisiGarp to communicate knowledge about system behaviour, and that VisiGarp can be used by people from various backgrounds in different academic settings.

In the future, more attention could be paid to embedding VisiGarp in an educational context, including illustrative examples and case-studies which could indicate the added value of simulation and explanation tools. This would require more domain-specific materials in addition to our current generic approach. Authoring tools such as SimQuest (van Joolingen and de Jong 2003) may prove useful in creating a complete simulation-based learning environment, including more educational context and instructional features.

Supporting learning with multiple representations

A flexible tool such as VisiGarp, with several different views, menu choices and buttons to choose from, carries the danger of the 'art museum' problem (Foss 1989)—students may be overwhelmed by the amount of information presented to them, and keep clicking on things without real understanding of what they are doing or what they are looking at. Integrating the information that is shown in different displays may also be an issue (Ainsworth 1999; Bodemer et al. 2005). Therefore, support is necessary to keep students on track and create a useful learning experience (van der Hulst 1996; Veermans 2002). In the VisiGarp evaluation study, this was done by providing on-screen exercises that the participants had to complete, organized in an order from structure to behaviour, and from relatively simple to more complex simulation data. These exercises were designed by hand, but it may be possible to use automatically generated tutoring exercises (Goddijn et al. 2003), if this is tied to a curriculum of didactic goals (Nuttle et al. 2005).

Conclusion

Qualitative simulation is a powerful tool for predicting and explaining system behaviour in various domains. The explicit representation of all kinds of knowledge involved makes it a fruitful resource for education. Communicating the information contained in qualitative simulations is not a trivial task, however. In this article, *diagrammatic visualization* has been investigated as the means of communication for this purpose. A tool, VisiGarp, has been developed, which generates diagrams of any simulation model expressed in the format of Garp, a framework for QR. The interface of VisiGarp is organized around a number of views, each of which focuses on particular kinds of information. There are views which show generic information in a domain library, and there are views which show the results of a specific simulation, such as the state–transition graph, the quantity value history view, and the dependencies view. The state–transition graph shows the qualitative states in which the system may manifest itself, and the possible order. The quantity value history view shows how the value and derivative of quantities in the simulation change over time. The dependencies view shows the details of the system in a particular state, including the entities, attributes and structural relations involved, but also the causal and mathematical dependencies which hold between quantities. The implementation of VisiGarp shows that diagrammatic representations can be generated automatically from a model represented in a QR terminology, thereby creating a link between conceptual knowledge and formal simulation models.

VisiGarp contributes to the state of the art in the form of the visualization principles behind the design of the visualizations, including the mapping from model ingredients to a set of visual primitives applicable to qualitative simulation. Furthermore, it is one of the few systems, next to the Expound system (Mallory 1998), capable of generating explanatory diagrams for qualitative multi-state simulations. Compared to this work, VisiGarp includes a larger set of views, showing a richer set of model ingredients. Furthermore, VisiGarp's diagrams are an integrated part of the interface, which supports the navigation process.

VisiGarp was designed to support interactive investigation of qualitative simulation models and results. To determine whether it fulfills this goal, an evaluation study was carried out with thirty students. The results of this study indicate that students without knowledge of QR quickly learned to use VisiGarp, and rated it positively afterwards. Furthermore, while using VisiGarp, the students were able to run qualitative simulations and answer questions about system behaviour in a domain unfamiliar to them, and they actually learned something about the domain in the process. These findings clearly show that the interactive visualizations provided by VisiGarp allow students to inspect qualitative simulations, and have potential for use in educational settings to support learning about the behaviour of (ecological) systems. VisiGarp is also considered a useful tool by students and domain experts who have built simulation models themselves. This is confirmed by the fact that VisiGarp is in actual use for research and education purposes at several universities.

Further work will address ways to incorporate textual explanations to complement the visualizations, to aggregate the information from simulations, and to integrate VisiGarp's functionality with tools for building models.

Acknowledgements The authors would like to thank Peggy Tjaris for her role in carrying out the evaluation studies and analysis of the results, Paulo Salles for allowing us to use his model of the Brazilian Cerrado ecology, and Bob Wielinga and Joost Breuker for their comments on earlier versions of this article.

Appendix: VisiGarp treatment exercise questions

This appendix lists the exercise questions used in the treatment during the evaluation of VisiGarp. They are meant to introduce novice students to the VisiGarp tool and the domain of population ecology, especially in the Brazilian Cerrado. Two simulation scenarios are used. Scenario 1 is about a single population. Scenario 2 is about the Cerrado, which consists of three different populations. The original treatment also contained instructions on how to use VisiGarp (left out for brevity). The original version of the questions was in Dutch (Tjaris 2002), aimed at the Dutch participants in the evaluation study. For convenience, they are translated here into English.

Scenario 1

1. Which of the following statements is true?
   a. open_population1 consists_of population1.
   b. *population1 consists_of biological_species1.*
   c. open_population1 is an assumption about biological_species1.
2. Which of the terms below is **not** a quantity (characteristic) of the entity **population1**?
   a. *biological_species1.*
   b. born1.
   c. number_of1.
3. Which values can the number_of the population adopt?
   a. zero, plus.
   b. *zero, normal, max.*
   c. zero, plus, normal, max.

4. Which values can the number of births (born) adopt?
   a. zero, normal, max.
   b. *zero, plus.*
   c. min, zero, plus.
5. Which value does the number_of the population have? And does this number increase or decrease in this state?
   a. number_of has the value zero and is increasing.
   b. number_of has the value normal and is decreasing.
   c. *number_of has the value normal and is increasing.*
6. Which value does born of the population have? Does born increase or decrease in this state?
   a. *born has the value plus and is increasing.*
   b. born has the value plus en is decreasing.
   c. born has the value zero and is decreasing.
7. Study the selected path. Which statement about this path is true?
   a. the number_of the population changes from state 8 to 1 from zero to max, and from state 2 to 6 in the reverse direction, i.e. from max to zero.
   b. *the number_of the population changes from state 8 to 1 from max to zero, and from state 2 to 6 in the reverse direction, i.e. from zero to max.*
   c. the number_of the population changes in the selected path from max to zero.
8. Why is there no increase of the number of the population yet in state 1, and why is there in state 2?
   a. There is less emigration in state 2.
   b. Born and dead are in balance in state 1, but not in state 2.
   c. *There is no colonization yet in state 1, but there is in state 2.*
9. Why does the growth of the population increase?
   a. *born is equal to dead AND immigration is greater than emigration.*
   b. born is equal to dead AND immigration is smaller than emigration.
   c. born is equal to dead AND immigration is equal to emigration.
10. There is no change in the growth of the population, because
    a. *born and dead are equal and immigration and emigration are equal.*
    b. born and dead are equal and immigration is greater than emigration.
    c. born is greater than dead and immigration is smaller than emigration.

Scenario 2

11. Which statement about the Cerrado vegetation is correct?
    a. the vegetation consists of 2 populations, that is, populations of grass and shrubs.

b. *the vegetation consists of 3 populations, that is, populations of grass, shrubs, and trees.*
c. the vegetation consists of 2 populations, that is, populations of shrubs and trees.

12. What is the task of the Cerrado manager within the presented simulation model?

a. to control fire, especially increasing the frequency of fires.
b. to merge different flows going to zero.
c. *to control fire, especially decreasing the frequency of fires.*

13. Which quantities are relevant for describing a population (for example, the tree population)?

a. number_of, dead, born, and fire_frequency.
b. *number_of, dead, born, and growth.*
c. number_of, inflow, outflow, growth, and cover.

14. Which values can the number_of a population (for example the grass population) adopt?

a. zero, plus.
b. *zero, low, medium high and max.*
c. min, zero and plus.

15. What is the value of cover of the Cerrado vegetation? And does cover increase or decrease in this state?

a. cover has the value plus and is decreasing.
b. *cover has the value low and is increasing.*
c. cover has the value plus and is increasing.

16. What is the value of number_of the tree population? And does number_of increase or decrease in this state?

a. number_of has the value plus and is decreasing.
b. *number_of has the value low and is increasing.*
c. number_of has the value high and is decreasing.

17. Study the selected path. Which statement about this path is correct?

a. the number_of the grass and shrub population changes from zero to max and from zero to high, respectively, while the number_of the tree population decreases from max to zero.
b. *the number_of the tree- and shrub population changes from zero to max and from zero to high, respectively, while the number_of the grass population decreases from max to zero.*
c. the number_of the grass and tree population changes from zero to max and from zero to high, respectively, while the number_of the shrub population decreases from max to zero.

18. Study how the values of number_of for the tree, shrub, and grass population compare over all states. Which statement best characterizes these changes over the different states?

a. the tree population decreases, mostly from max to zero or low, while the grass and shrub population mostly increase from zero to max or high.
b. the shrub population decreases, mostly from max to zero, while the grass and tree population increase, mostly from zero to max or high.
c. *the grass population decreases, mostly from max to zero, while the shrub and tree population increase, mostly from zero to max or high.*

19. What kind of dependency is there between fire frequency and litter in the Cerrado vegetation?

a. When fire frequency decreases, litter also decreases.
b. *When fire frequency decreases, litter increases.*
c. There is no direct dependency between these two quantities.

20. What kind of dependency is there between litter and soil_temperature in the Cerrado vegetation?

a. When litter increases, soil_temperature also increases.
b. *When litter increases, soil_temperature decreases.*
c. There is no direct dependency between these two quantities.

21. What kind of dependency is there between cover and litter in the Cerrado vegetation?

a. *When cover increases, litter also increases.*
b. When cover increases, litter decreases.
c. There is no direct dependency between these two quantities.

22. What kind of dependency is there in the Cerrado vegetation between litter and the factors soil_temperature, light, moisture and nutrient?

a. Changes in litter (e.g. increase) propagate to the same changes in light and soil_temperature (they also increase), and changes in the opposite direction for moisture and nutrient (they decrease).
b. *Changes in litter (e.g. increase) propagate to the same changes in moisture and nutrient (they also increase), and changes in the opposite direction for light and soil_temperature (they decrease).*
c. Changes in litter (bijv. increase) propagate to the same changes in light and moisture (they also increase), and changes in the opposite direction for soil_temperature and nutrient (they decrease).

23. What can be deduced about the growth of the tree population?

a. *the tree population increases, because immigration is equal to emigration, while born is greater than dead.*
b. the tree population increases, because immigration is greater than emigration, while born is equal to dead.
c. the tree population increases, because immigration is greater than emigration, AND born is greater than dead.

24. What can be deduced about the growth of the shrub population?

a. *the shrub population increases, because immigration is equal to emigration, while born is greater than dead.*

b. the shrub population increases, because immigration is greater than emigration, while born is equal to dead.
c. the shrub population increases, because immigration is greater than emigration, AND born is greater than dead.

25. The factors soil_temperature, light, moisture, and nutrient have an effect on born and dead of the populations in the Cerrado vegetation. Which of the following statements about the tree population is correct?

    a. When litter increases, the composition of soil_temperature, light, moisture and nutrient changes in such a way that for the tree population, born (growth of new trees) becomes equal to dead (dying of existing trees).
    b. *When litter increases, the composition of soil_temperature, light, moisture and nutrient changes in such a way that for the tree population, born (growth of new trees) becomes greater than dead (dying of existing trees).*
    c. When litter increases, the composition of soil_temperature, light, moisture and nutrient changes in such a way that for the tree population, born (growth of new trees) becomes smaller than dead (dying of existing trees).

26. The factors of soil_temperature, light, moisture and nutrient have an effect on born and dead of the populations in the Cerrado vegetation. Which of the following statements about the grass population is correct?

    a. When litter increases, the composition of soil_temperature, light, moisture and nutrient in such a way that for the grass population born, (growth of new grass) becomes equal to dead (dying of existing grass).
    b. When litter increases, the composition of soil_temperature, light, moisture and nutrient changes in such a way that for the grass population, born (growth of new grass) becomes greater than dead (dying of existing grass).
    c. *When litter increases, the composition of soil_temperature, light, moisture and nutrient in such a way that for the grass population, born (growth of new grass) becomes smaller than dead (dying of existing grass).*

27. In this simulation a manager regulates the fire frequency in the Cerrado vegetation. Which of the following statements is correct? (Note: Try to deduce the answer from the *Dependencies* shown in the *Dependencies* view)

    a. When the manager succeeds in decreasing the fire frequency, the grass and shrub population will increase, while the tree population will decrease.
    b. When the manager succeeds in decreasing the fire frequency, the grass and tree population will increase, while the shrub population will decrease.
    c. *When the manager succeeds in decreasing the fire frequency, the tree and shrub population will increase, while the grass population will decrease.*

References

Ainsworth, S. (1999). The functions of multiple representations. *Computers and Education, 33*(2–3), 131–152.

Alvarez-Bravo, J. V., Alvarez-Sanchez, J. J., & Gonzalez-Cabrera, F. J. (2004). Learning physical concepts using a qualitative approach: A teaching proposal. In J. de Kleer & K. D. Forbus (Eds.), *Proceedings of*

QR 2004, the 18th International Workshop on Qualitative Reasoning, August 2–4, 2004. Evanston, IL: Northwestern University.

Bodemer, D., Ploetzner, R., Bruchmuller, K., & Hacker, S. (2005). Supporting learning with interactive multimedia through active integration of representations. *Instructional Science, 33*(1), 73–95.

Bouwer, A. (2005). *Explaining behaviour: Using qualitative simulation in interactive learning environments*. PhD Thesis, Universiteit van Amsterdam, July 2005.

Bouwer, A., Machado, V. B., & Bredeweg, B. (2002). Interactive model building environments. In P. Brna, M. Baker, K. Stenning, & A. Tiberghien (Eds.), *The role of communication in learning to model* (Chap. 6, pp. 155–182). London: Lawrence Erlbaum Associates.

Bredeweg, B. (1992). *Expertise in qualitative prediction of behaviour*. PhD Thesis, University of Amsterdam, Amsterdam, The Netherlands.

Bredeweg, B., Bouwer, A., Jellema, J., Bertels, D., Linnebank, F., & Liem, J. (2006a). Garp3: A new workbench for qualitative reasoning and modelling. In F. Wotawa (Eds.), *Proceedings of the 3rd Model Based Systems Workshop, part of the 17th European Conference on Artificial Intelligence (ECAI 2006), 28 August–1st September* (pp. 67–74). Italy: Riva del Garda.

Bredeweg, B., Salles, P., & Neumann, M. (2006b). Ecological applications of qualitative reasoning. In F. Recknagel (Ed.), *Ecological informatics: Scope, techniques and applications* (2nd ed., pp. 15–47). Berlin: Springer.

Bredeweg, B., & Struss, P. (2003). Current topics in qualitative reasoning (editorial introduction). *AI Magazine, 24*(4), 13–16.

Bredeweg, B., & Winkels, R. (1998). Qualitative models in interactive learning environments: An introduction. *Interactive Learning Environments, 5*, 1–18.

Cañas, A. J., Carff, R., Hill, G., Carvalho, M., Arguedas, M., Eskridge, T. C., Lott, J., & Carvajal, R. (2005). Concept maps: Integrating knowledge and information visualization. In S.-O. Tergan & T. Keller (Eds.), *Knowledge and information visualization: Searching for synergies* (pp. 343–354). Heidelberg: Springer Lecture Notes in Computer Science.

Cheng, P. C.-H., Lowe, R. K., & Scaife, M. (2001). Cognitive science approaches to understanding diagrammatic representations. *Artificial Intelligent Review, 15*(1–2), 79–94.

de Kleer, J., & Brown, J. (1984). A qualitative physics based on confluences. *Artificial Intelligence, 24*, 7–83.

de Kleer, J. H. (1990) Qualitative pysics: A personal view. In D. S. Weld & J. H. de Kleer (Eds.), *Readings in qualitative reasoning about physical systems* (pp. 1–8). San Mateo, CA: Morgan Kaufmann.

de Koning, K., Bredeweg, B., Breuker, J., & Wielinga, B. (2000). Model-based reasoning about learner behaviour. *Journal of Artificial Intelligence, 117*(2), 173–229.

Doerr, H. M. (1996). STELLA ten years later. *International Journal of Computers for Mathematical Learning, 1*(2), 201–224.

Elio, R., & Sharf, P. B. (1990). Modeling novice-to-expert shifts in problem-solving and knowledge organization. *Cognitive Science, 14*, 579–639.

Engelhardt, Y. (2002). *The Language of Graphics: A framework for the analysis of syntax and meaning in maps, charts and diagrams*. PhD Thesis, University of Amsterdam.

Forbus, K. D. (1984). Qualitative process theory. *Artificial Intelligence, 24*, 85–168.

Forbus, K. D. (1988). Qualitative physics: Past, present and future. In H. E. Shrobe (Ed.), *Exploring artificial intelligence* (pp. 239–296). San Mateo, CA: Morgan Kaufmann.

Forbus, K. D., & Feltovich, P. J. (Eds.). (2001). *Smart machines in education*. Cambridge, MA: AAAI Press/MIT Press.

Forbus, K., Carney, K., Harris, R., & Sherin, B. (2001). A qualitative modeling environment for middle-school students: A progress report. In G. Biswas (Ed.), *Proceedings of QR 2001, 15th International Workshop on Qualitative Reasoning, St. Mary's University, San Antonio, Texas, 17–18 May 2001* (pp. 65–72). Stoughton, WI: The Printing House.

Forbus, K. D., Whalley, P. B., Everett, J. O., Ureel, L., Brokowski, M., Baher, J., & Kuehne, S. E. (1999). Cyclepad: An articulate virtual laboratory for engineering thermodynamics. *Artificial Intelligence, 114*(1/2), 297–347.

Forrester, J. (1968). *Principles of systems*. Cambridge, MA: MIT Press.

Foss, C. L. (1989). Detecting lost users: Empirical studies on browsing hypertext. Technical Report, Computer-Based Learning Unit, University of Leeds.

Frederiksen, J. R., & White, B. Y. (2002). Conceptualizing and constructing linked models: Creating coherence in complex knowledge systems. In P. Brna, M. Baker, K. Stenning, & A. Tiberghien (Eds.), *The role of communication in learning to model* (pp. 69–96). London: Lawrence Erlbaum Associates.

Goddijn, F., Bouwer, A., & Bredeweg, B. (2003). Automatically generating tutoring questions for qualitative simulations. In P. Salles & B. Bredeweg (Eds.), *Proceedings of the 17th International workshop on Qualitative Reasoning, QR'03* (pp. 87–94). Brazil: Brasilia.

Grimm, V. (1994). Mathematical models and understanding in ecology. *Ecological Modelling, 75/76*, 641–651.

Harel, D. (1995). On visual formalisms. In J. Glasgow, N. H. Narayanan, & B. Chandrasekaran (Eds.), *Diagrammatic reasoning* (pp. 235–271). Cambridge, MA: MIT Press.

Hollan, J. D., Hutchins, E. L., & Weizenbaum, L. (1984). STEAMER: an interactive inspectable simulation-based training system. *AI Magazine, 5*(2), 15–27.

Jackson, S., Stratford, S., Krajcik, J., & Soloway, E. (1996). Making system dynamics modeling accessible to pre-college science students. *Interactive Learning Environments, 4*(3), 233–257.

Kuipers, B. J. (1994). *Qualitative reasoning: Modeling and simulation with incomplete knowledge*. Cambridge, MA: MIT Press.

Kulpa, Z. (1994). Diagrammatic representation and reasoning. *Machine GRAPHICS & VISION, 3*(1/2), 77–103.

Larkin, J., & Simon, H. (1987). Why a diagram is (sometimes) worth ten thousand words. *Cognitive science, 11*, 65–99.

Leelawong, K., Wang, Y., Biswas, G., Vye, N., Bransford, J., & Schwartz, D. (2001). Qualitative reasoning techniques to support learning by teaching. In G. Biswas (Ed.), *Proceedings of QR 2001, 15th International Workshop on Qualitative Reasoning, St. Mary's University, San Antonio, Texas, 17–18 May 2001* (pp. 65–72). Stoughton, WI: The Printing House.

Mallory, R. S. (1998). *Tools for explaining complex qualitative simulations*. PhD Thesis, Department of Computer Sciences, University of Texas at Austin.

Mettes, C. T. C. W., & Roossink, H. J. (1981). Linking factual and procedural knowledge in solving science problems: A case study in a thermodynamics course. *Instructional Science, 10*, 333–361.

Neumann, M., & Bredeweg, B. (2004). A qualitative model of the nutrient spiraling in lotic ecosystems to support decision makers for river management. In J. de Kleer & K. D. Forbus (Eds.), *Proceedings of QR 2004, the 18th International Workshop on Qualitative Reasoning, August 2–4*. Evanston, IL: Northwestern University.

Norman, D. A. (1993). *Things that make us Smart: Defending human attributes in the age of the machine*. Cambridge, MA: Perseus Books.

Novak, J. D., & Gowin, D. B. (1984). *Learning how to learn*. New York: Cambridge University Press.

Nuttle, T., Bredeweg, B., & Salles, P. (2004). Qualitative reasoning about food webs: Exploring alternative representations. In J. de Kleer & K. D. Forbus (Eds.), *Proceedings of QR 2004, the 18th International Workshop on Qualitative Reasoning, August 2–4*. Evanston, IL: Northwestern University.

Nuttle, T., Salles, P., & Bredeweg, B. (2005). Guidelines for sustainable development curriculum, project no. 004074, project deliverable report D6.8. Technical Report, Naturnet-Redime, STREP Project co-funded by the European Commission within the Sixth Framework Programme (2002–2006).

Pivello, V. R. (1992) *An expert system for the use of prescribed fires in the Management of Brazilian Savannas*. PhD Thesis, Imperial College, London.

Ploetzner, R., & Spada, H. (1998). Constructing quantitative problem representations on the basis of qualiative reasoning. *Interactive Learning Environments, 5*, 95–108.

Reichherzer, T. R., Cañas, A. J., Ford, K. M., & Hayes, P. J. (1998). The Giant: A classroom collaborator. In C. Frasson, G. Gouardères, L. W. Johnson, J. Lester, & J. Rickel (Eds.), *Workshop Notes of the ITS '98 Workshop on Pedagogical Agents*, San Antonio, TX.

Salles, P. (1997). *Qualitative models in ecology and their use in learning environments*. PhD Thesis, University of Edinburgh, Scotland, UK.

Salles, P., & Bredeweg, B. (1997). Building qualitative models in ecology. In *Proceedings of the International workshop on Qualitative Reasoning, QR'97*, pp. 155–164, Italy, June 1997, Istituto di Analisi Numerica C.N.R. Pavia.

Salles, P., & Bredeweg, B. (2001). Constructing progressive learning routes through qualitative simulation models in ecology. In G. Biswas (Ed.), *Proceedings of the 15th International workshop on Qualitative Reasoning, QR'01, May 17–19* (pp. 82–89). San Antonio, TX: IOS-Press.

Salles, P., & Bredeweg, B. (2003a) A case study of collaborative modelling: Building qualitative models in ecology. In U. Hoppe, F. Verdejo, & J. Kay (Eds.), *Artificial intelligence in education: Shaping the future of learning through intelligent technologies* (pp. 245–252). Osaka, Japan: IOS-Press/Ohmsha.

Salles, P., & Bredeweg, B. (2003b). Qualitative reasoning about population and community ecology. *AI Magazine, 24*(4), 77–90.

Salles, P., & Bredeweg, B. (2006). Modelling population and community dynamics with qualitative reasoning. *Ecological Modelling, 195*(12), 114–128.

Salles, P., Bredeweg, B., Araujo, S., & Neto, W. (2003). Qualitative models of interactions between two populations. *AI Communications, 16*(4, 24), 291–308.

Salles, P., Gauche, R., & Virmond, P. (2004). A qualitative model of the daniell cell for chemical education. In J. de Kleer & K. D. Forbus (Eds.), *Proceedings of QR 2004, the 18th International Workshop on Qualitative Reasoning, August 2–4, 2004*. Evanston, IL: Northwestern University.

Soloway, E., Pryor, A. Z., Krajcik, J. S., Jackson, S., Stratford, S. J., Wisnudel, M., & Klein, J. T. (1997). Scienceware's model-it: Technology to support authentic science inquiry. *Technological Horizons on Education, 25*(3), 54–56.

Suthers, D. D. (2003). Representational guidance for collaborative learning. artificial intelligence in education. In H. U. Hoppe, F. Verdejo, & J. Kay (Eds.), *Proceedings of AI-ED 2003, the 11th International Conference on Artificial Intelligence in Education* (pp. 3–10). Amsterdam: IOS Press, Keynote address.

Tjaris, P. (2002) Kwalitatieve simulaties als middel tot kennisoverdracht. Master's Thesis, Social Science Informatics (SWI), Universiteit van Amsterdam. In Dutch.

Tullos, D. D., Neumann, M., & Alvarez-Sanchez, J. J. (2004). Development of a qualitative model for investigating benthic community response to anthropogenic activities. In J. de Kleer & K. D. Forbus(Eds.), *Proceedings of QR 2004, the 18th International Workshop on Qualitative Reasoning, August 2–4*. Evanston, IL: Northwestern University.

Tversky, B. (1995) Cognitive origins of graphic conventions. In F. T. Marchese (Ed.), *Understanding images* (pp. 29–53). New York: Springer-Verlag.

van Berkum, J. J. A., & de Jong, T. (1991). Instructional environments for simulations. *Journal of Education & Computing, 6*(3/4), 305–358.

van der Hulst, A. (1996). *Cognitive tools: Two exercises in non-directive support for exploratory learning*. PhD Thesis, University of Amsterdam, Amsterdam.

van Joolingen, W. R. (1999). Cognitive tools for discovery learning. *International Journal of Artificial Inteligence and Education, 10*, 385–397.

van Joolingen, W. R., & de Jong, T. (2003). SimQuest, authoring educational simulations. In *Authoring Tools for Advanced Technology Learning Environments: Toward cost-effective adaptive, interactive, and intelligent educational software* (pp. 1–31). Dordrecht: Kluwer.

van Joolingen, W. R., de Jong, T., & Dimitrakopoulou, A. (2007). Issues in computer supported inquiry learning in science. *Journal of Computer Assisted Learning, 23*(2), 111–119.

Veermans, K. (2002). *Intelligent support for discovery learning: Using opportunistic learner modeling and heuristics to support simulation based discovery learning*. PhD Thesis, University of Twente.

Weld, D. S., & de Kleer, J. H. (1990). *Readings in qualitative reasoning about physical systems*. San Mateo, CA: Morgan Kaufmann.